

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-244695

(43)Date of publication of application : 19.09.1997

(51)Int.Cl. G10L 9/00

G10L 9/14

// H03M 7/30

(21)Application number : 08-046191

(71)Applicant : KOBE STEEL LTD

(22)Date of filing : 04.03.1996

(72)Inventor : NISHIMOTO YOSHIRO

TAKAHASHI TETSUYA

SHIMODA TOSHIAKI

SAKATANI TORU

HIEKATA TAKAYUKI

(54) VOICE CODING DEVICE AND DECODING DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide encoding and decoding techniques which can carry out encoding and decoding without a grating noise based on a background noise when users communicate with each other with cellular phones and so on.

SOLUTION: A transmitter (encoder) smoothes in time a fluctuation of feature quantities of voice, in only a silence section. On the other hand, a receiver (decoder) smoothes in time a fluctuation of feature quantities of voice, in only a silence section, and replaces a part of driving signal with other signal, so that the grating noise is suppressed.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-244695

(43) 公開日 平成9年(1997)9月19日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/00			G 1 0 L 9/00	H
				C
				F
9/14			9/14	Q
// H 0 3 M 7/30		9382-5K	H 0 3 M 7/30	B
審査請求 未請求 請求項の数 8 O L (全 13 頁)				

(21) 出願番号 特願平8-46191

(22) 出願日 平成8年(1996)3月4日

(71) 出願人 000001199

株式会社神戸製鋼所

兵庫県神戸市中央区脇浜町1丁目3番18号

(72) 発明者 西元 善郎

兵庫県神戸市西区高塚台1丁目5番5号

株式会社神戸製鋼所神戸総合技術研究所内

(72) 発明者 高橋 哲也

兵庫県神戸市西区高塚台1丁目5番5号

株式会社神戸製鋼所神戸総合技術研究所内

(72) 発明者 下田 敏章

兵庫県神戸市西区高塚台1丁目5番5号

株式会社神戸製鋼所神戸総合技術研究所内

(74) 代理人 弁理士 本庄 武男

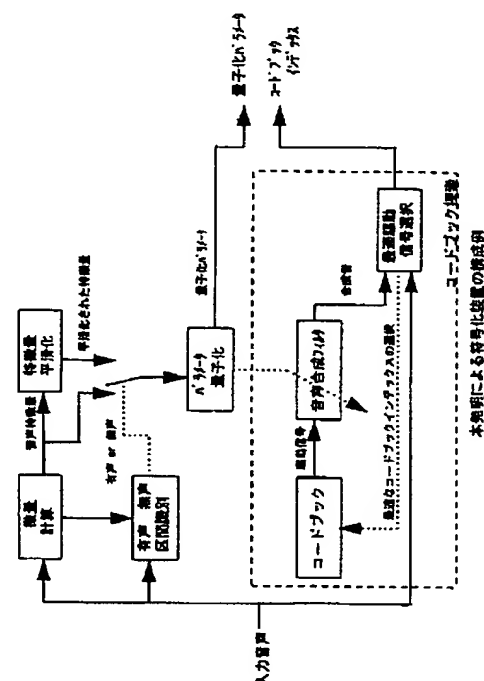
最終頁に続く

(54) 【発明の名称】 音声符号化装置及び復号化装置

(57) 【要約】

【課題】 携帯電話等の電話通信時における背景音による耳障りな雑音のない音声の符号化及び復号化を達成すること。

【解決手段】 音声信号からその特徴量を計算し、そのパラメータを圧縮（符号化）すると共に、受信側で原音に近い音声信号を合成するための音声合成用駆動信号に対応するインデックスを選択し、上記圧縮されたパラメータとインデックスを送信すると共に、上記パラメータとインデックスを受信した受信側では、上記パラメータとインデックスから音声特徴量と駆動信号を復元し、音声合成する音声符号装置及び復号化装置において、送信側（符号化時）においては、音声の無音区間における音声の特徴量の変動を時間的に平滑化すると共に、受信側（復号化時）においては音声の無音区間における音声の特徴量の変動を時間的に平滑化するとともに駆動信号の一部を他の信号に置きかえることにより、背景音を符号及び復号化した時の上記耳障りな雑音を抑制するようにした符号化装置及び復号化装置。



【特許請求の範囲】

【請求項 1】 音声信号から音声特徴量を計算し、上記音声特徴量を構成するパラメータを圧縮して送信すると共に、上記音声信号に応じて音声合成用の駆動信号に対応するインデックスをコードブックから選択して上記圧縮されたパラメータと共に送出する音声符号化装置において、上記音声信号内における有声・無声区間を識別する符号化有声・無声区間識別手段と、上記無声区間内においてのみ上記音声特徴量の変動を時間的に平滑化するか、無声区間においては有声区間よりも強い平滑化を行うかのいずれかの平滑化処理を行う符号化平滑化手段とを具備してなることを特徴とする音声符号化装置。

【請求項 2】 上記音声特徴量として線スペクトル対を計算し、無声区間では各線スペクトル対の値を時間的に平滑化した後に音声特徴量を量子化する請求項 1 記載の音声符号化装置。

【請求項 3】 音声符号化装置からの出力である圧縮されたパラメータ及び駆動信号生成のためのインデックス等を受信し、これらのパラメータ及びインデックス等から音声特徴量及び駆動信号を復元し、それらを用いて音声合成する音声復号装置において、上記音声符号化装置からの圧縮されたパラメータから上記音声信号内の有声・無声部分を識別する復号化有声・無声区間識別手段と、上記無声区間内においてのみ上記音声特徴量の変動を時間的に平滑化するか、無声区間においては有声区間よりも強い平滑化を行うかのいずれかの平滑化処理を行う復号化平滑化手段と、平滑化した後の各音声特徴量を用いて音声合成する音声合成手段とを具備してなることを特徴とする音声復号化装置。

【請求項 4】 上記音声特徴量として線スペクトル対を計算し、無声区間では各線スペクトル対の値を時間的に平滑化した後に上記駆動信号を計算し、音声合成する請求項 3 記載の音声復号化装置。

【請求項 5】 音声符号化装置からの出力である圧縮されたパラメータ及び駆動信号生成のためのインデックス等を受信し、これらのパラメータ及びインデックス等から音声特徴量及び駆動信号を復元し、それらを用いて音声合成する音声復号装置において、符号化装置からからの圧縮されたパラメータから上記音声信号内の有声・無声部分を識別する復号化有声・無声区間識別手段と、無声区間と判断された部分では、上記インデックスを元に上記コードブックから復元された第 1 の駆動信号の一部若しくは全てを他の方法で生成された第 2 の駆動信号で置き換えた第 3 の駆動信号を作成し、この第 3 の駆動信号に基づいて音声合成することを特徴とする音声復号化装置。

【請求項 6】 第 2 の駆動信号として乱数等によって発生される雑音信号を用い、第 3 の駆動信号をコードブックから復元された第 1 の駆動信号と第 2 の駆動信号の重み付け加算により作成する請求項 5 記載の音声復号化装

置。

【請求項 7】 有声・無声の切り替わり境界部では、上記第 3 の駆動信号を作成する際に重み付け加算する第 1 の駆動信号と乱数等の第 2 の駆動信号との割合を連続的に変化させる請求項 5 もしくは請求項 6 のいずれかに記載の音声復号化装置。

【請求項 8】 無声部分では音声特徴量の変動を時間的に平滑化し、平滑化した後の特徴量と上記駆動信号を用いて音声合成する請求項 5、請求項 6 若しくは請求項 7 のいずれか 1 つに記載の音声復号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、携帯電話等で問題となる背景雑音に基づく異音を抑制するための音声符号化装置及び復号化装置に関するものである。

【0002】

【従来の技術】本発明の適用対象となる音声圧縮・復元（符号化・復号化）装置の一例として、CELP (Code Excited Linear Prediction) に基づく携帯電話の基本構成は、図 10 及び図 11 に示されている。図 10 は音声符号化装置、図 11 は音声復号化装置の構成例を示す。図 10 に示す符号化装置は、入力された原音声から音声の特徴量を計算する「特徴量計算部」と、それらを量子化してデータ圧縮する「パラメータ量子化部」と、後に音声合成するために用いる駆動信号に対応するインデックスを複数記憶している「コードブック部」と、原音と近い音を再現するための最適なコードブック内のインデックスを選択する「最適駆動信号選択部」とを具備して構成されている。上記最適駆動信号選択部では、合成される音声は原音声に近い音声になるような駆動信号に対応するインデックスが上記コードブックの中の複数のインデックスの内から選択される。即ち、最適駆動信号選択部は、入力音声を受信すると、コードブック内のインデックスから複数のインデックスを選択し、これを順次音声合成フィルタに入力して合成音を作成する手順を選択したインデックスの数だけ繰り返し、選択されたインデックスの数の合成音の中から入力音声に最も近い合成音を生成するインデックスを抽出して出力する。この時、音声合成フィルタは、前記音声特徴量に対応する係数に応じて駆動される。

【0003】符号化されたデータを復号化するには、量子化された音声特徴量や選ばれたインデックス等が図 11 の復号化装置に送信され、「音声特徴量復号部」において特徴量の値が復号化され、「駆動信号生成部」では、符号化装置から送信されたインデックスに基づいてコードブックからの駆動信号が作成される。また、音声のスペクトル包絡形状を表す特徴量を用いて、「音声合成フィルタ」の係数が決められ、そのフィルタに駆動信号を入力することにより音声合成される。ここで、駆動信号は人間の声帯の振動等の音源に対応し、スペクト

3

ル包絡形状を作りだす合成フィルタは声道の形状等により音源から発せられるまでの伝達関数に対応する。従って、音声特徴量を量子化する際の量子化テーブル並びにコードブックに格納されているインデックスは、人間の音声を合成するのに適したものが、種々の規格に基づいて用意されている。

【0004】その一方で、このような方式で低ビットレートに音声を圧縮する音声圧縮・復元方式では、人間の声以外の音が入力された場合には、必ずしも元の音を忠実に再現することが出来ず、特に空調音や雨音等、ごく身近に存在する背景雑音が混在した場合でも、極めて音質が劣化することが知られている。その一つはswirling noiseと呼ばれるものであり、背景雑音が「キュルキュル」いったふうに変動する極めて不快な音になるという現象である。この原因は、スペクトル包絡形状を表す音声特徴量等を計算するのに通常使われるフレームでは、雑音の特徴量を安定して計算するためにはデータが少なすぎ、音声パラメータが変動するためであることが、論文“Improvement of Background Sound coding in Linear Predictive Speech Codes”, Proc. ICASSP 1995 等に明らかにされている。

【0005】

【発明が解決しようとする問題点】このような問題を解決するための従来技術の一つとしては、特開平7-152395号公報に記載のように、無声区間において背景雑音のスペクトル特徴量を求め、それに基づいて雑音スペクトルを抑圧するフィルタにより雑音を減衰させる方式がある。しかし、この方式ではS/N比は改良されるが、雑音を完全に無くすことは不可能であり、残った雑音から計算される特徴量が変動して不快に聞こえるため、問題の本質は解決できていない。また、電話等においては背景雑音も周囲の臨場感を伝える情報であり、背景の音がなくなってしまうことにより、かえって不自然な印象を与えるという別の問題も発生する。

【0006】また別の従来技術として特開平7-160294号公報では、復号化装置側で雑音のコードブックを用意しておき、符号化装置から送信されるインデックス等によって生成される信号の代わりに、無声区間では「雑音コードブック」から選んだ別の信号を使う。この場合「雑音コードブック」からの信号を、合成フィルタに入力する前の駆動信号と置き換える方法と、合成フィルタの出力である合成信号と置き換える方法とが示されているが、そのいずれもが問題を有している。まず、合成フィルタ前の信号を変えても、合成フィルタの係数はスペクトル包絡形状の音声特徴量からきめられるため、swirling noiseの原因が音声パラメータの変動である以上、結局はフィルタから出力される合成音声には不快な変動が生じることになる。逆に、合成フィルタ後の音声を「雑音コードブック」からの信号に置き換える方法では、存在する可能性のあるスペクトル特性を持った雑音信号を

4

全てコードブックとして用意しておかなければ、原音を再現できないという問題がある。さらに、コードブックの探索は、CELP方式において最も計算量を必要とする部分であり、符号化装置だけでなく復号化装置においてもコードブック探索を行うことは、極めて高速な信号処理用演算装置が必要となってしまう。

【0007】また、別の従来技術として特開平7-036485号公報では、過去の音声信号を記憶しておくバッファを設けておき、無声区間では通常の音声区間の計算よりも多いフレーム数のデータを用いて音声特徴量を計算することにより、計算結果を安定にして特徴量の変動を抑えようとしている。しかし、前述の論文によれば、計算結果を安定させるために320ms程度の音声データを使用しており、8KHz サンプリングのデジタル音声データでは、2560 サンプルものデータが必要となる。従って、この方法では、バッファだけのために多くのメモリ資源が必要となってしまう。従って本発明の目的は、大きいメモリ容量を必要とすることなく、また高速の演算装置を用いることなく、前記背景雑音に基づく不快音を生じることのない音声符号化装置及び復号化装置を提供することである。

【0008】

【課題を解決するための手段】上記目的を達成するために本願が採用する音声符号化装置は、音声信号から音声特徴量を計算し、上記音声特徴量を構成するパラメータを圧縮して送信すると共に、上記音声信号に応じて音声合成用の駆動信号に対応するインデックスをコードブックから選択して上記圧縮されたパラメータと共に送出する音声符号化装置において、上記音声信号内における有声・無声区間を識別する符号化有声・無声区間識別手段と、上記無声区間内においてのみ上記音声特徴量の変動を時間的に平滑化するか、無声区間においては有声区間よりも強い平滑化を行うかのいずれかの平滑化処理を行う符号化平滑化手段とを具備してなることを特徴とする音声符号化装置として構成されている。この場合、上記音声特徴量としては、線スペクトル対を計算し、無声区間では各線スペクトル対の値を時間的に平滑化した後に音声特徴量を量子化するものが挙げられる。

【0009】また本願における音声復号化装置に関する第1発明は、音声符号化装置からの出力である圧縮されたパラメータ及び駆動信号生成のためのインデックス等を受信し、これらのパラメータ及びインデックス等から音声特徴量及び駆動信号を復元し、それらを用いて音声を合成する音声復号装置において、上記音声符号化装置からの圧縮されたパラメータから上記音声信号内の有声・無声部分を識別する復号化有声・無声区間識別手段と、上記無声区間内においてのみ上記音声特徴量の変動を時間的に平滑化するか、無声区間においては有声区間よりも強い平滑化を行うかのいずれかの平滑化処理を行う復号化平滑化手段と、平滑化した後の各音声特徴量を

5

用いて音声合成する音声合成手段とを具備してなることを特徴とする音声復号化装置として構成されている。この場合、上記音声特徴量として線スペクトル対を計算し、無声区間では各線スペクトル対の値を時間的に平滑化した後に上記駆動信号を計算し、音声合成することができる。

【0010】また第2発明は、音声符号化装置からの出力である圧縮されたパラメータ及び駆動信号生成のためのインデックス等を受信し、これらのパラメータ及びインデックス等から音声特徴量及び駆動信号を復元し、それらを用いて音声合成する音声復号化装置において、符号化装置からからの圧縮されたパラメータから上記音声信号内の有声・無声部分を識別する復号化有声・無声区間識別手段と、無声区間と判断された部分では、上記インデックスを元に上記コードブックから復元された第1の駆動信号の一部若しくは全てを他の方法で生成された第2の駆動信号で置き換えた第3の駆動信号を作成し、この第3の駆動信号に基づいて音声合成することを特徴とする音声復号化装置として構成されている。ここに上記第2の駆動信号を、乱数等によって発生される雑音信号を用い、第3の駆動信号をコードブックから復元された第1の駆動信号と第2の駆動信号の重み付け加算により作成することができる。また、有声・無声の切り替わり境界部では、上記第3の駆動信号を作成する際に、重み付け加算する第1の駆動信号と乱数等の第2の駆動信号との割合を連続的に変化させるようにしても良い。さらに、上記無声部分では、音声特徴量の変動を時間的に平滑化し、平滑化した後の特徴量と上記駆動信号を用いて音声合成するようにしても良い。

【0011】

【発明の実施の形態】続いて、添付図面を参照して本発明を具体化した実施例につき説明し、本発明の理解に供する。尚、以下に示す実施例は本発明を具体化した一例であって、本発明の技術的範囲を限定するものではない。ここに、図1は、本発明の一実施形態に係る音声符号化装置の一構成を示すブロック図、図2は、本発明の一実施形態に係る音声復号化装置の一構成を示すブロック図、図3は、特徴量に関する線スペクトル対の変動を示すグラフ、図4は、図3に示した線スペクトル対の量子化後の変動を示すグラフ（平滑化なし）、図5は、図3に示した線スペクトル対の量子化後の変動を示すグラフ（平滑化あり）、図6は、本発明の一実施例に係る音声復号化装置の構成を示すブロック図、図7は、上記図6に示した音声復号化装置における特徴量のパラメータの平滑化を行う構成を示すブロック図、図8は、原音声の背景雑音のスペクトルを示すグラフ、図9は、合成された背景雑音のスペクトルを示すグラフ、図10は、従来のCELP符号化装置の構成を示すブロック図、図11は、従来のCELP復号化装置の構成を示すブロック図である。

6

【0012】図1および図2は、本発明の一実施形態に係る音声符号化装置及び音声復号化装置の構成を示す。まず音声符号化装置（図1）では入力された音声信号から「特徴量計算部」にて音声の特徴量となる各種パラメータが計算される。次に「有声・無声区間識別部」において上記特徴量に基づいて、入力信号が現時点で音声を含んだ有声区間か、背景雑音のみの無声区間かを判断する。無声区間と判断された場合には、計算された特徴量は「特徴量平滑化部」において変動を抑制される。逆に、有声区間と判断された場合には、特徴量の平滑化は行わずに計算値をそのまま使う。これ以降は、前記従来技術で説明した図10のCELP方式の符号化装置等と同様であり、特徴量は「パラメータ量子化部」において量子化（圧縮）される。また音声が入力されると、最適駆動信号選択部においてコードブック部を検索し、複数のインデックスの中から、原音に近い音を再現するための最適なインデックスを「コードブック部」から検索する。こうして符号化装置の出力である量子化された特徴量パラメータと、コードブック探索結果であるインデックスが復号化装置に送信される。

【0013】復号化装置では、図2に示すように、それらに基づいて音声合成を行う。すなわち、まず「音声特徴量復号部」では、量子化されているパラメータを音声特徴量の値に逆変換する。次に、「有声・無声区間識別部」では、特徴量に基づいて、音声の状態が有声区間か、無声区間かを判断する。もちろん、符号化装置の「有声・無声区間識別部」で識別された結果を、陽にパラメータとして含んでいてもよい。無声区間と判断された場合には、計算された特徴量は「特徴量平滑化部」において変動を抑制される。逆に、有声区間と判断された場合には、特徴量はの平滑化は行わずに計算値をそのまま使う。このような平滑化処理を受けた音声のスペクトル包絡形状の特徴量から、音声合成するための音声合成フィルタのパラメータが決定される。また、合成のための入力に使う駆動信号もインデックスから復元され、前記特徴量パラメータにより決定される「合成フィルタ部」に駆動信号を入力することにより、原音に近い音声合成される。多くの場合、音声の復号化装置は、音声合成フィルタ部の後に「ポストフィルタ部」（不図示）を有し、合成された音声の聴覚上の品質を高めて出力される。

【0014】続いて各部分の具体的計算手法を示すが、一般的に用いられる音声信号処理のアルゴリズムの説明に関しては「音声情報処理の基礎（オーム社）」等の参考書に記載されており、一般的なCELP方式と同様な処理部分に関しては「RCR規格27C」等の規格書等に詳しく記載されているので、ここでは省略する。「特徴量計算部」で計算される主なパラメータには、音声の音量、スペクトル包絡形状、ピッチ等がある。特徴量の計算は、フレームと呼ばれる時間的に連続した音声デー

7

タ $X(n)$, $n=1, 2, \dots, N_A$ を用いた分析によって行う。先ず音量のパラメータとしては、次式のようなフレーム内での音声信号の平均振幅 A 等が採用可能である。

【0015】

【数1】

$$A = (1/N_A) \sqrt{\sum_{n=1}^{N_A} X^2(n)}$$

$$\begin{pmatrix} R_0 & R_1 & R_2 & \dots & R_{M-1} \\ R_1 & R_2 & R_3 & \dots & R_{M-2} \\ R_2 & R_1 & R_2 & \dots & R_{M-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{M-1} & R_{M-2} & R_{M-3} & \dots & R_0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_M \end{pmatrix} + \begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_M \end{pmatrix} = 0$$

ここで、 R_i , $i=0, 1, 2, \dots, M$ は連続した音声データ $X(n)$, $n=1, 2, \dots, N_A$ を用いて次式により計算した自己相関である。ただし、通常ここで R_i の計算に用いられる N_A 個の音声データは、ハミングウィンドウなどを施されたものが使われる。

【0017】

【数3】

$$R_i = \sum_{n=1}^{N_A-i} X(n) \cdot X(n+i)$$

この方程式を解く方法としては、Durbinの再帰的解法等を用いれば効率的であることが、よく知られている。また、ピッチを求める手法としては、自己相関法等がある。下式の Z 変換で表されるような逆合成フィルタに音声 $X(n)$ を通した出力として得られる予測残差信号の自己相関関数を計算すれば、音声のピッチに対応する部分で自己相関関数がピークを持つことが知られている。

【0018】

【数4】

$$1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} \dots + \alpha_M z^{-M}$$

ピッチに関してはこのような計算方法以外に、後で述べるコードブック探索において駆動信号が最適な周期性を持つように評価関数で選択する方法もよく知られている。「有声/無声区間識別」において無声部を検出する手段としては、音声に基づく方法や信号の周期性に着目した方法などが用いられる。つまり、比較的小さい音量で変動が少ないフレームが続いた場合に、無声区間と判定する方法や、有声部分ではピッチによる周期性を持った信号になるということから周期性の少ない部分を無声区間と判定する方法などが知られている。周期性を表わ

8

また、スペクトル包絡形状の特徴量としては、下の方程式を解いて得られる線形予測係数 α_i , $i=1, 2, \dots, M$ がある。

【0016】

【数2】

すパラメータとしては、特徴量の計算で求められた線形予測係数による予測残差信号の自己相関関数のピーク値などがよく使われる。

【0019】「特徴量平滑化部」においては、特に背景雑音で生じる不快な音の変動に大きく寄与しているスペクトル包絡パラメータを時間的に平滑化する。ただし、上述の線形予測係数をそのまま平滑化した場合には、平滑化後のパラメータによって構成される合成フィルタが安定である保証はない。合成フィルタが不安定になると、合成される出力信号は発振音になってしまい、極めて不都合である。この問題を解決する方法として、線形予測係数 α_i , $i=1, 2, \dots, M$ を線スペクトル対 ω_i , $i=1, 2, \dots, M$ と呼ばれる等価なパラメータに変換した後に平滑化すればよい。線スペクトル対への変換は M が偶数であれば、 α_i から決まる下式の多項式の根を求めることで行える。

$$P(z) = 1 + (\alpha_1 + \alpha_{10})z^{-1} + (\alpha_2 + \alpha_{10})z^{-2} \dots + (\alpha_{10} + \alpha_1)z^{-10} + z^{-11}$$

$$Q(z) = 1 + (\alpha_1 - \alpha_{10})z^{-1} + (\alpha_2 - \alpha_{10})z^{-2} \dots + (\alpha_{10} - \alpha_1)z^{-10} + z^{-11}$$

【0020】この2個の多項式 $P(z)$ 及び $Q(z)$ の根は安定であれば、 $z_i = e^{-j\omega_i}$ の形の根を持ち、このときの $P(z)$ の根 $0 < \omega_1 < \omega_3 < \omega_5 < \dots < \omega_{M-1} < \pi$ および $0 < \omega_2 < \omega_4 < \omega_6 < \dots < \omega_M < \pi$ が線スペクトル対と呼ばれる。線スペクトル対は $0 < \omega_i < \omega_j < \pi$ ($i < j$)という条件を満たす限りは安定であることが知られているので、元のパラメータが安定であれば平滑化後も容易に安定性を維持できる。例えば、第 k フレームの線スペクトル対を $\omega_i(k)$, $i=1, 2, \dots, M$ とすれば、次式のように平滑化を行うことができる。

9

$$\Omega_i(k) = \beta \cdot \Omega_i(k-1) + (1-\beta) \cdot \omega_i(k)$$

(ただし、初期状態として $\Omega_i(1) = \omega_i(1)$ とする)

$0 < \beta < 1$ であり、 β が大きいほど平滑化は強く行われる。明らかに、平滑化前の線スペクトル対 $\omega_i(k)$ が上記の安定条件を満たせば、平滑化後の $\Omega_i(k)$ も安定条件を満たす。

【0021】無声区間においては、このようにして得られる平滑化された線スペクトル対 $\Omega_i(k)$ 、 $i=1, 10$

$$(1/2)(P(z) + Q(z)) = 1 + a_1 z^{-1} + a_2 z^{-2} \dots + a_M z^{-M}$$

有声区間においてパラメータを平滑化してしまうと、人の声のスペクトルを歪めてしまうことになるので平滑化せずに元の ω_i を用いるか、もしくは声が歪まない程度で雑音によるパラメータ変動を軽減させる目的で β の小さな軽い平滑化を行ってもよい。

【0023】平滑化の効果は、スペクトル包絡形状を表わす特徴量に対して施すのが最も効果的であるが、もちろん音量やピッチなどの他の特徴量に対して行ってもよい。「パラメータ量子化」「コードブック」「コードブック探索」に関しては、各種の規格において様々なものが用いられており、本実施例でもそれらに準拠したものを使うことができる。例えば、「パラメータ量子化部」ではスカラー量子化や「ベクトル量子化」によってパラメータのビット数を削減し、データの圧縮がされる。また、通常「コードブック」は、過去の駆動信号を記憶しておき周期波形を作り出す適応コードブックと、周期波形だけでは表わしきれない部分を補償するための固定的なコードブックから成るのが一般的である。

【0024】このコードブックから生成可能な駆動信号のうち、入力の実音に近い音声を合成する最適な駆動信号を選択するのが「コードブック探索」である。ここでは、聴覚重み付けされた歪みを表わす評価関数が最小となるような波形を選択する方法が通常用いられている。その際に、周期波形を作り出す適応コードブックに対しては過去に選ばれた駆動信号から周期波形を作り出すためのピッチやゲインが決定されインデックスとして出力され、固定的な信号波形をテーブルとして持っているコードブック部では、それらの信号の中から選ばれた最適な波形の組み合わせや符号・ゲインなどがインデックスとして出力される。符号化装置から出力された量子化後のパラメータやコードブックのインデックスは、復号化装置への入力となり、「特徴量復号部」において特徴量の値が復元される。また、コードブックのインデックスからは符号化装置で選択された駆動信号が生成される。復元された特徴量は、符号化装置と同様に復号化装置においても平滑化することが可能である。これにより、データ送信元の符号化装置が特徴量平滑化の機能を持たないものであっても、復号化装置において平滑化による効果

10

2, ..., Mにもとづいて合成フィルタを構成すれば、背景雑音部における音質の不快な変動が軽減される。平滑化後の線スペクトル対に対応する合成フィルタの係数は、 $e^{-j \cdot Q_i}$ を根に持つ多項式を展開して作り出せば上記の多項式 $P(z)$ 、 $Q(z)$ が構成できることから、 α_i に対応する量子化後の線形予測係数 a_i 、 $i=1, 2, \dots, M$ を以下のように求めることができる。

【0022】

【数5】

を与えることができる。「合成フィルタ」の特性は、複号スペクトル包絡形状を表わす特徴量によって設定される。本実施例では、平滑化された線スペクトル対を線形予測係数に変換した a_i 、 $i=1, 2, \dots, M$ を用い、次式の z 変換で表わされるフィルタに駆動信号を入力することにより音声を合成する。

【0025】

【数6】

$$\frac{1}{1 + a_1 z^{-1} + a_2 z^{-2} \dots + a_M z^{-M}}$$

このように合成された音声は、より聴覚的な音質を向上させるために通常なんらかの「ポストフィルタ」で処理する。ポストフィルタとしては、高域強調、スペクトル整形、ピッチ強調などのフィルタがよく用いられる。

【0026】以上述べた実施の形態では復号化装置に、符号化装置から送られてきた音声パラメータなどに基づき無声区間/有声区間を判定する機能を持たせ、無性と判断された区間においては符号化装置から送られてきたインデックスなどによってコードブックから生成される駆動信号をそのまま使用しているので、コードブックから生成される駆動信号が人の声に適したものだけであり、背景雑音に対して使用した際に不自然な音となる。これをより自然な雑音に聞こえる音に変えることができる実施例を次に示す。この際、有声部分と無声部分とで背景雑音の質が急に変化する不自然感を緩和するため、境界部分ではコードブックから復元した第1の駆動信号と乱数などの他の方法で発生させた第2の駆動信号から重み付け加算などによって第3の駆動信号を作り出す際に、両者を加え合わせる重み付けを連続的に変化させてもよい。さらに、無声部分において、符号化装置から送られてきた音声特徴量を時間的に平滑化して背景雑音に起因して発生する特徴量の変動を抑える方法と、上記の駆動信号の加工処理とを併用することにより、背景雑音の不快感をさらに軽減できる。

【0027】

【実施例】まず、符号化装置からは元の音声から計算された音声特徴量を量子化したパラメータや、コードブッ

クから駆動信号を生成するためのインデックスが出力され、復号化装置への入力となる。復号化装置「特徴量復号」部において特徴量の値が復元される。また、コードブックのインデックスからは符号化装置で選択された駆動信号が生成される。復号化装置の場合には、図 6 に示す如く、符号化装置から送られてきた量子化された音量のパラメータから音量による有声・無声区間の判断が可能である。また、コードブックから生成される駆動信号は人の声帯などで発生するピッチ周期の振動に対応しており、この駆動信号の自己相関関数から周期性を検出することが可能であり、有声・無声の判定に用いることができる。

【0028】「駆動信号生成部」では符号化装置で選ばれたコードブックのインデックスに基づいて「コードブック」部から第 1 の駆動信号が生成される。また、本実施例では乱数などによる「ランダム雑音発生部」を設

$$g1 = \sqrt{1 - K}$$

$$g2 = \sqrt{K \cdot \frac{\sum e_1^2(j)}{\sum e_2^2(j)}}$$

もちろん、 $g1$ と $g2$ の値をこれよりも少なめにし、無声部における背景雑音を抑えぎみに出力することもできる。有声部では、基本的に第 1 の駆動信号をそのまま ($K=0$) で使用するが、このような駆動信号の置き換えによる有声部と無声部の境界での背景雑音の質の急激な変化を軽減するために、有声部と無声部の切り換えの際に K を連続的に変化させることも可能である。生成された駆動信号は、「合成フィルタ」を通すことにより音声

【0030】

【数 8】

$$\frac{1}{1 + a_1 z^{-1} + a_2 z^{-2} \dots + a_M z^{-M}}$$

合成フィルタに使われるスペクトル包絡を表わすパラメータは、符号化装置において元の音声から計算され量子化されて復号化装置に渡されるが、その際に量子化誤差の影響を受けにくくするなどの目的で、PARCOR 係数や線スペクトル対といったようなパラメータに変換された後に量子化するものが多い。これらは、線形予測係数と等価なパラメータなので、互いに変換することが可能である。これらの音声信号処理の基礎に関しては、参

け、第 1 の駆動信号とは異なり、より通常の背景雑音として自然な音源波形により第 2 の駆動信号を作り出す。無声部分と識別された場合には、「駆動信号加工部」において第 1 の駆動信号の全部もしくは一部を第 2 の駆動信号と置き換えて第 3 の駆動信号を作り出す。具体例としては、第 1 の駆動信号を $e1(i)$ 、第 2 の駆動信号を $e2(i)$ 、 $i=1, 2, \dots, N$ (N はフレームのサイズ) とすれば、以下のような重み付き線形和で第 3 の駆動信号を作ることができる。

$$e3(i) = g1 \cdot e1(i) + g2 \cdot e2(i)$$

$g1$ と $g2$ は $e1$ と $e2$ を加え合わせるための重みであり、例えば $e1$ と $e2$ が独立で相関がないとすると、次式のようにすれば元の駆動信号 $e1$ と同じ程度のパワーを持った $e3$ を作ることができる。

【0029】

【数 7】

(ただし、 $0 < K < 1$)

考等 (「音声情報処理の基礎」, オーム社などを参照) に詳しく記載されているので、ここでは詳細説明を省く。

【0031】特に、線スペクトル対の場合、時間的に平滑化しても安定性が保証されるので、無声部分などにおいてパラメータを平滑化することによって、背景雑音の不自然な変動が抑制され、さらに不快感を軽減させることも可能である。図 7 には、そのようなパラメータの平滑化も行う実施例が示してある。合成された音声は、より聴覚的な音質を向上させるために通常なんらかの「ポストフィルタ」(不図示) で処理する。ポストフィルタとしては、高域強調、スペクトル整形、ピッチ強調などのフィルタがよく用いられる。(RCR 規格 27C などを参照)

【0032】

【発明の効果】本発明によって得られる効果を以下に示す。また、図 3 には 20dB 程度の雨音が背景雑音として存在している入力音声から計算された線スペクトル対の値が示されている。線スペクトル対は、フレームの前後のデータも含めて 20ms (160 サンプル) のデータを使って計算した。この値に対して、平滑化を行わない従来の方式で線スペクトル対を量子化すると、例えば図 4 のようになる。これに対し、本発明によれば、符号化装置ならびに復号化装置において線スペクトル対は平滑化され、図 5 のようになる。これにより、無声区間における背景雑音部の特徴量変動が抑制されていることが分かり、聴覚による官能試験でも、音質の不自然な変動

による不快感が軽減されている。

【0033】無声部分における合成フィルタの特性は、背景雑音スペクトル特性を表わすことになる。したがって、e3を入力として音声を合成した場合にも、出力のスペクトルは図8、図9でみるように、元の背景雑音と類似したものとなる。図8と図9は、背景雑音である空調騒音のスペクトルと、本発明の複合化装置によって合成された後の音のスペクトルである。もちろん、この方式によって合成された背景雑音は元の音とは異なるものとなるが、官能試験によれば、第1の駆動信号をそのまま使う従来の方法によって出力される音よりも、第3の駆動信号によって合成された音のほうが、より自然で不快感がないという結果が得られた。

【図面の簡単な説明】

【図1】 本発明の実施の形態にかかる符号化装置の構成を示すブロック図。

【図2】 本発明の実施の形態にかかる複号化装置の構成を示すブロック図。

【図3】 線スペクトル対（特徴量計算部での値）の変動を示すグラフ。

【図4】 量子化後の線スペクトル対（平滑化しなかった場合の値）の変動を示すグラフ。

【図5】 量子化後の線スペクトル対（平滑化した値）の変動を示すグラフ。

【図6】 本発明の一実施例に係る音声復号化装置の構成を示すブロック図。

【図7】 図6に示した音声復号化装置における特徴量のパラメータの平滑化を行う構成を示すブロック図。

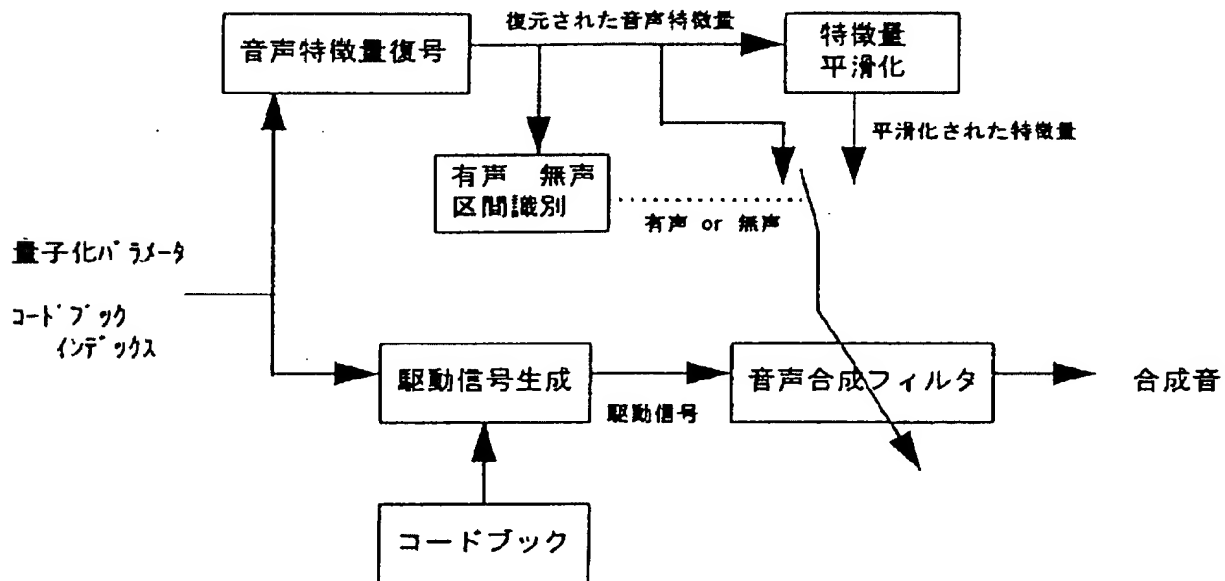
【図8】 原音声の背景雑音のスペクトルを示すグラフ。

【図9】 合成された背景雑音のスペクトルを示すグラフ。

【図10】 従来のCELP符号化装置の構成を示すブロック図。

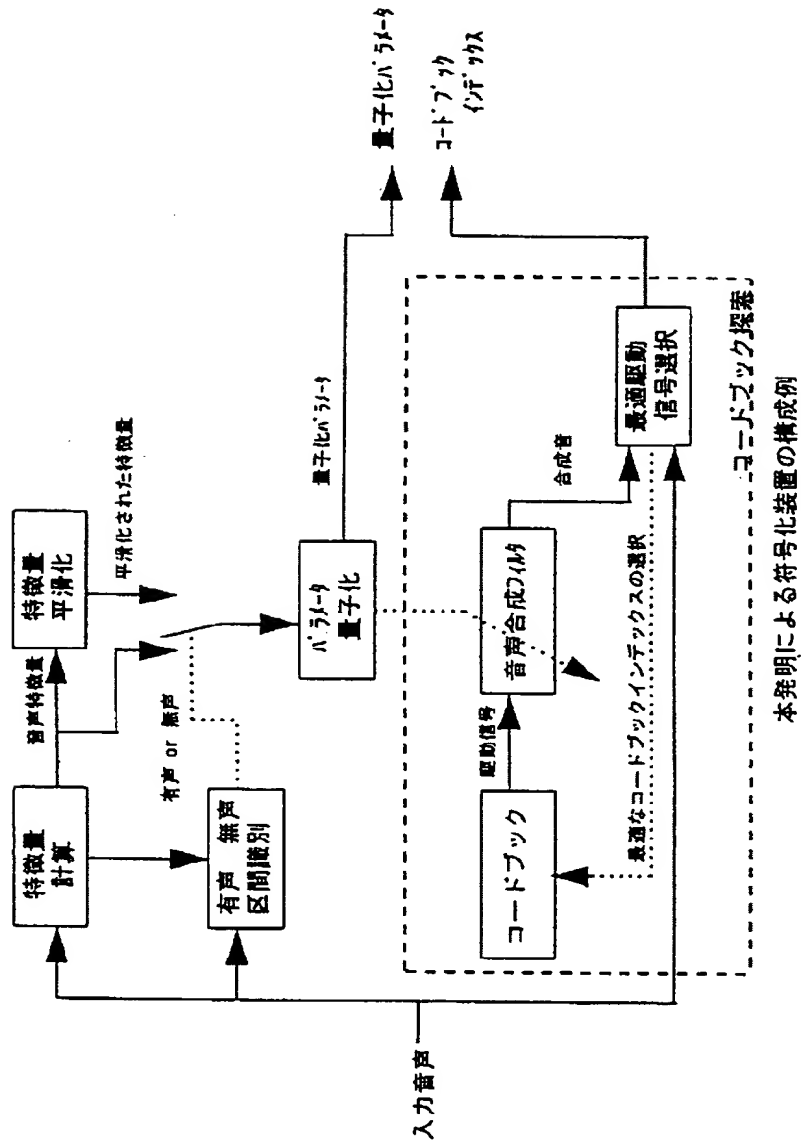
【図11】 従来のCELP復号化装置の構成を示すブロック図。

【図2】

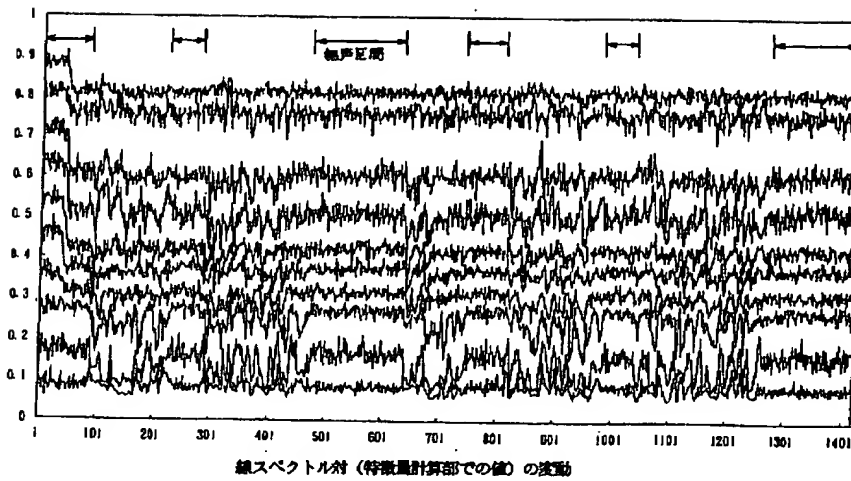


本発明による復号装置の構成

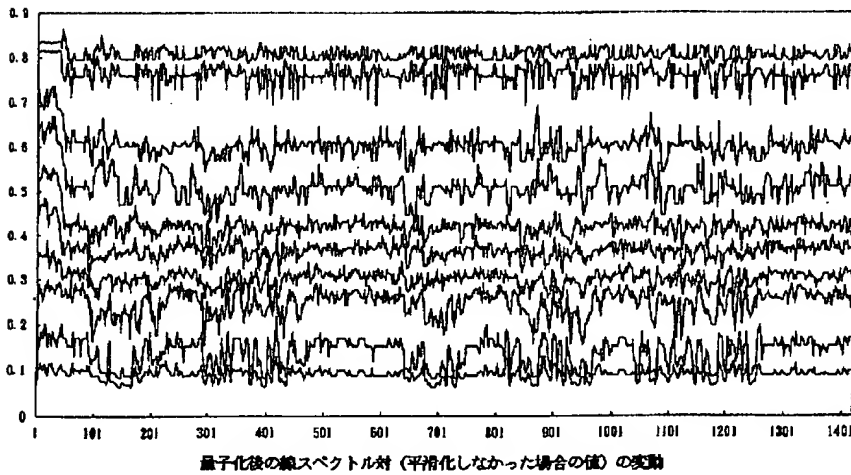
【図1】



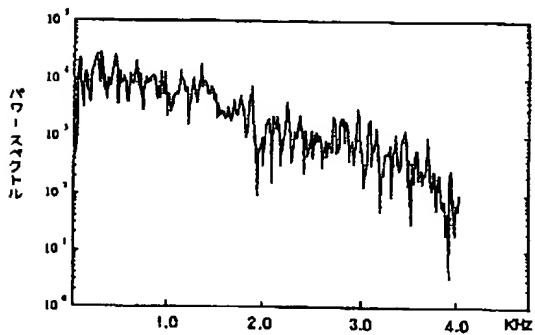
【図3】



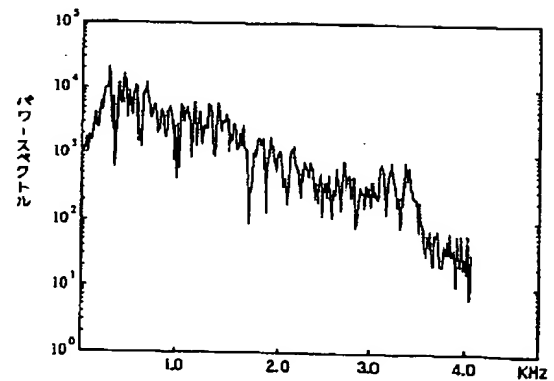
【図4】



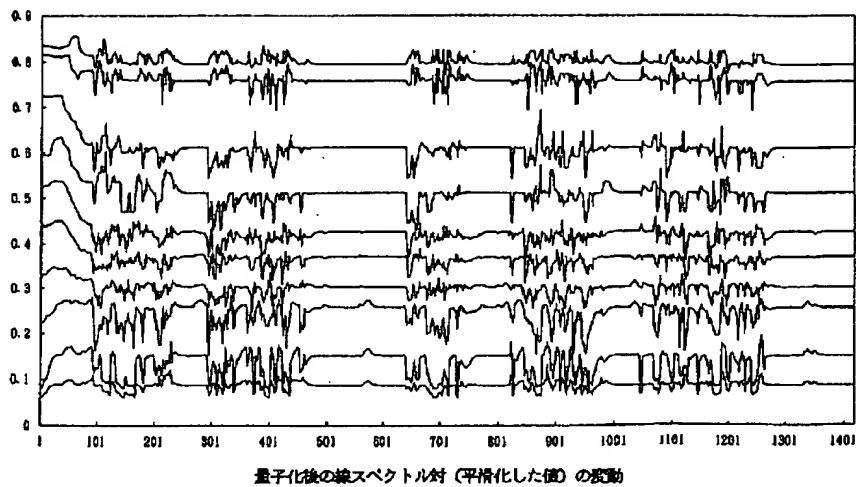
【図8】



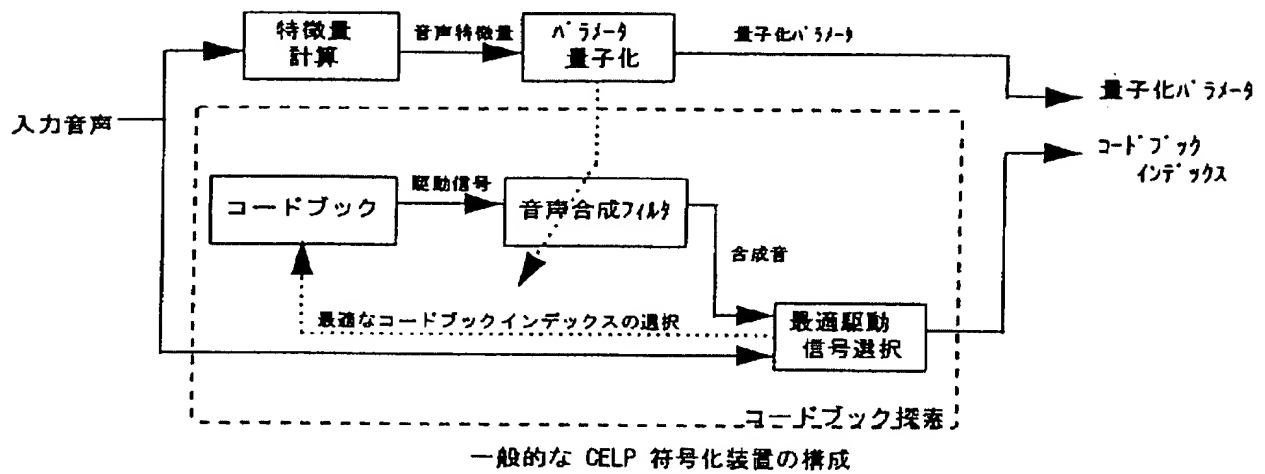
【図9】



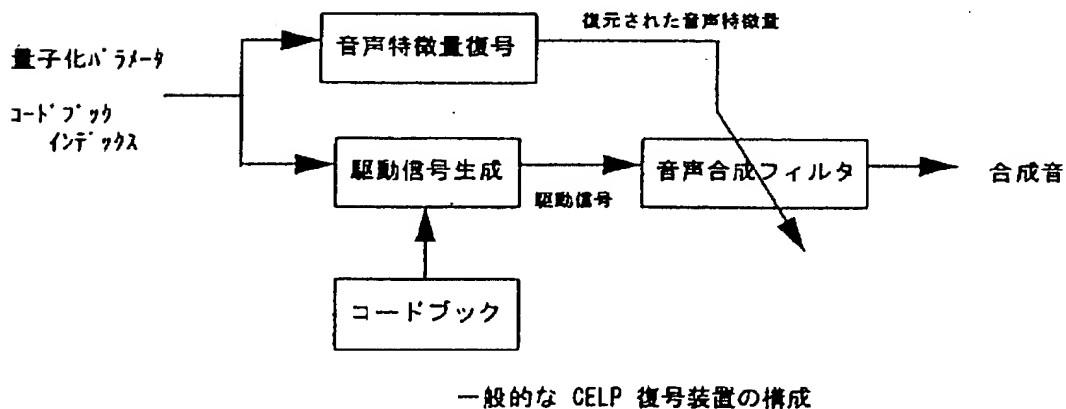
【図 5】



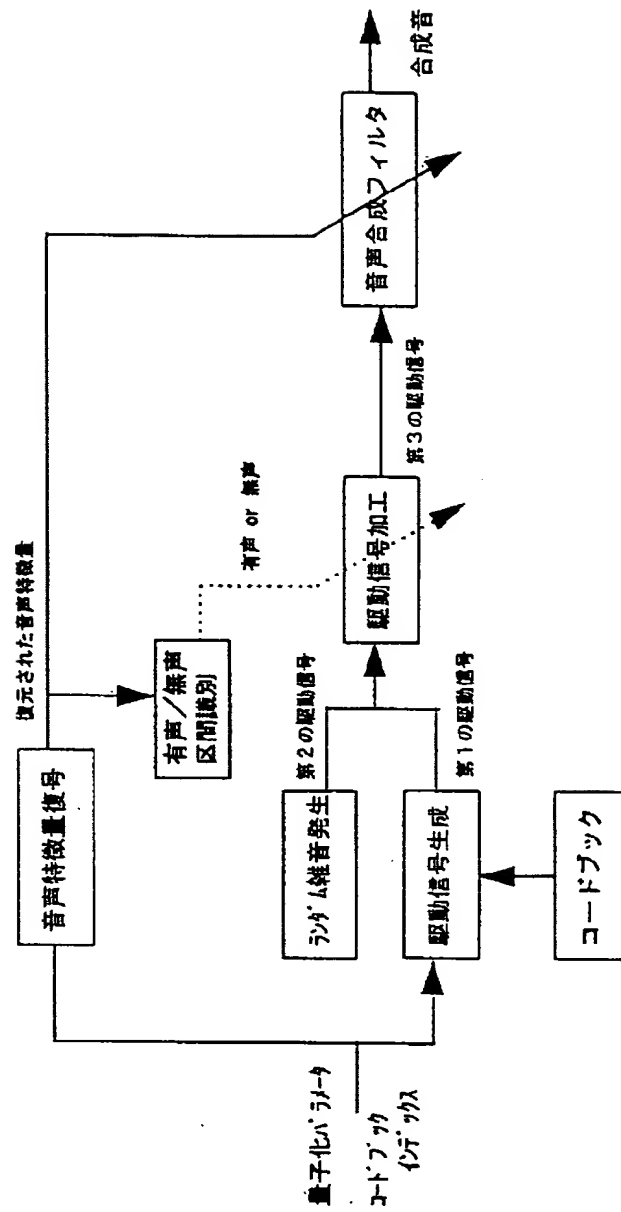
【図 10】



【図 11】

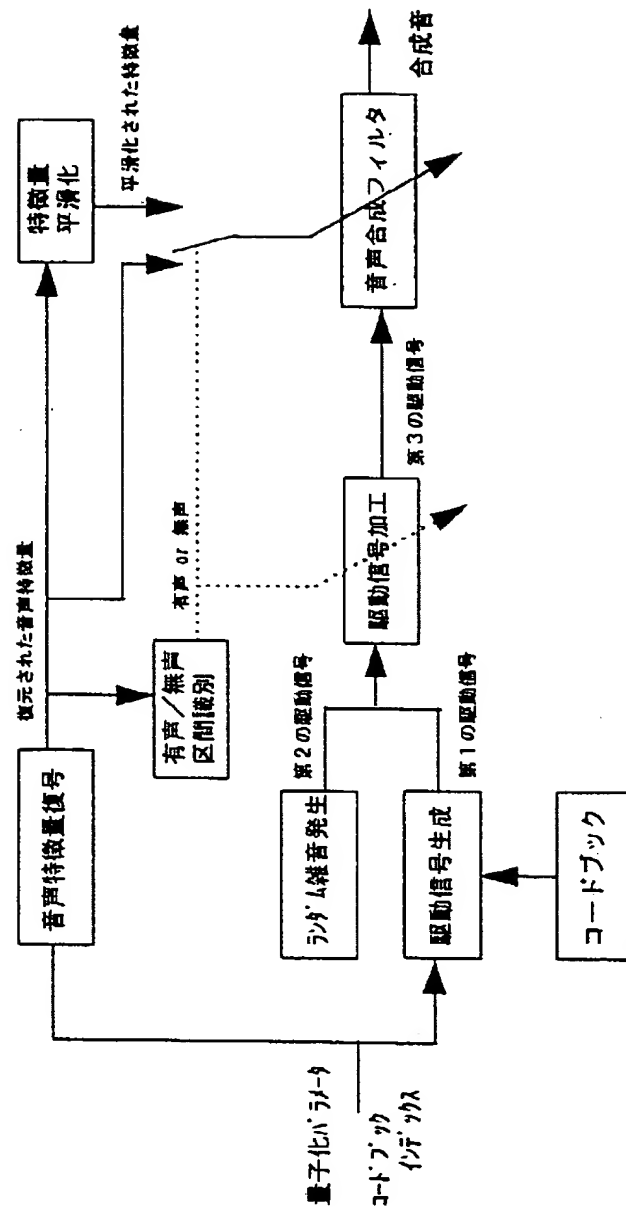


【図 6】



実施例の復号化装置

【図 7】



パラメータ平滑化も行う実施例

フロントページの続き

(72)発明者 坂谷 亨
 兵庫県神戸市西区高塚台1丁目5番5号
 株式会社神戸製鋼所神戸総合技術研究所内

(72)発明者 稗方 孝之
 兵庫県神戸市西区高塚台1丁目5番5号
 株式会社神戸製鋼所神戸総合技術研究所内